

# A Review of “Spikes not slots: Noise in neural populations limits working memory” by Bays (2015)

By Richard Thripp

EXP 6506 – University of Central Florida

September 10, 2015

This is an **opinion** article.

The author cites sources to create a case  
for his argument.

However, inferences are made that might  
not be made in a typical literature review.

# What is the slot model?

The idea that visual working memory (herein referred to as “**VWM**”\*) consists of 3–4 “slots” that can only represent a single visual object (p. 431).

\* Bays uses “WM” as his abbreviation, but I prefer “VWM” as a **constant reminder** that we are talking about **visual** working memory rather than working memory in general. Luck & Vogel (2013) use “VWM” as their abbreviation.



Image source: Super Mario 64 (1996 video game) [“select file” screen](#).

# What are spikes?

- Spikes are the firing of neurons.
- Their timing is **probabilistic**, roughly like the Poisson distribution.
- Recalling a VWM item requires enough spikes in the correct neurons (p. 432).

# **Deterministic Mechanism / Limit**

- A “fixed maximum number of representations that can be held in memory at one time” (p. 431).
- Or: Hard limit, ceiling, upper bound
- Encompasses the slot model and similar models.

# Implications of the Deterministic Model

- Represents a “hard limit” on VWM objects
- If more items must be remembered than slots available, some must be discarded



# Implications of the Deterministic Model

- Recall accuracy should have an “abrupt discontinuity” (p. 432) when the deterministic limit is exceeded.
- *However, Bays presents evidence that this abrupt discontinuity does not exist.*

# **“Stochastic”**

“Randomly determined; having a random probability distribution or pattern that may be analyzed statistically but may not be predicted precisely.”

SOURCE: [Oxford Dictionary](#) (U.S. English)

# **Stochastic Mechanism / Model**

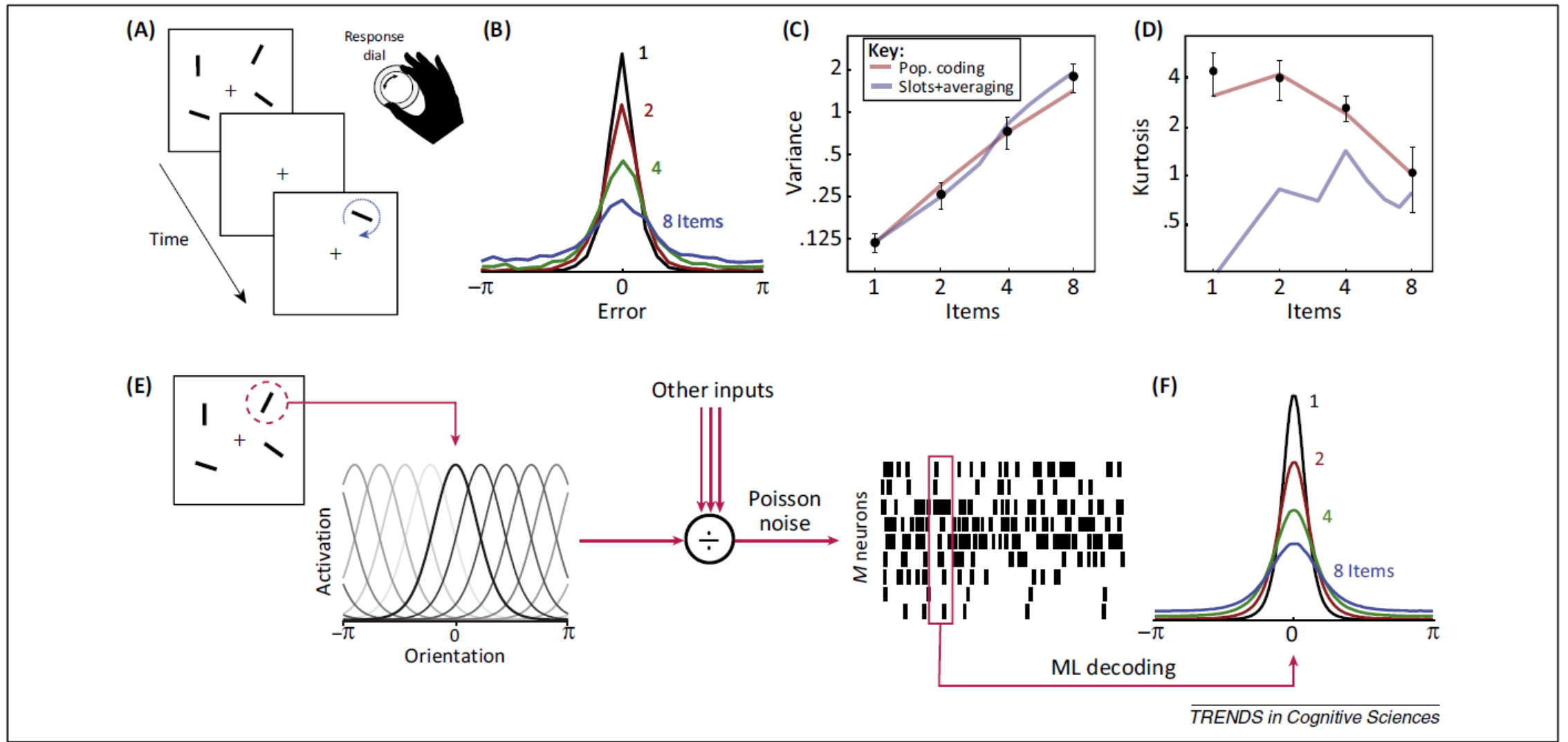
## **Or: Resource Model, Continuous Model**

“Representations in memory becoming increasingly variable as their number increases,” until they approach random noise (p. 431).



Image source: Wikipedia / public domain:  
[http://en.wikipedia.org/wiki/File:TV\\_noise.jpg](http://en.wikipedia.org/wiki/File:TV_noise.jpg)

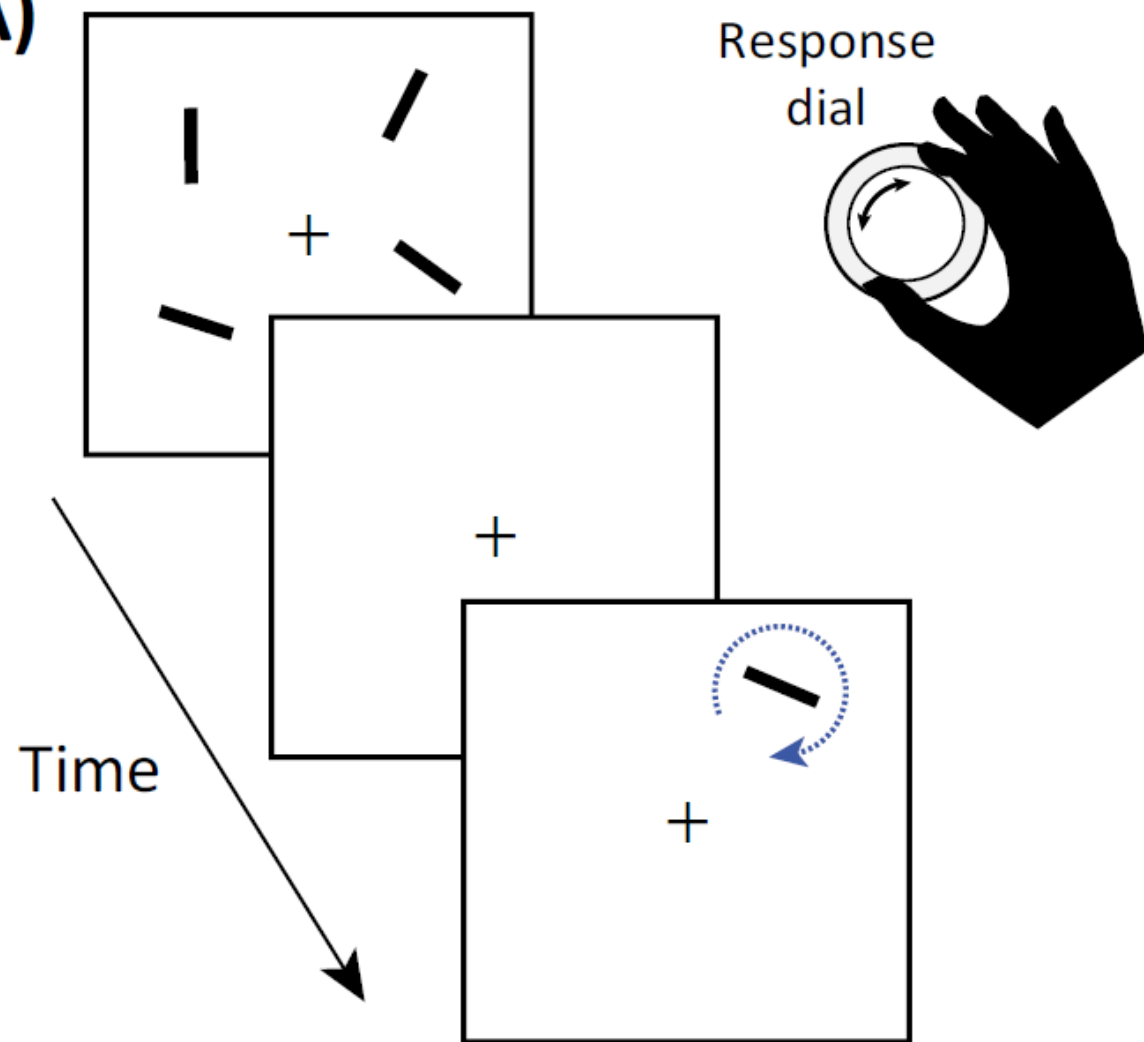
Key data for Bays' argument comes from analog recall tasks, where the subject must give a **continuous** (not multiple choice) response, such as turning a dial or selecting a color off a color wheel.



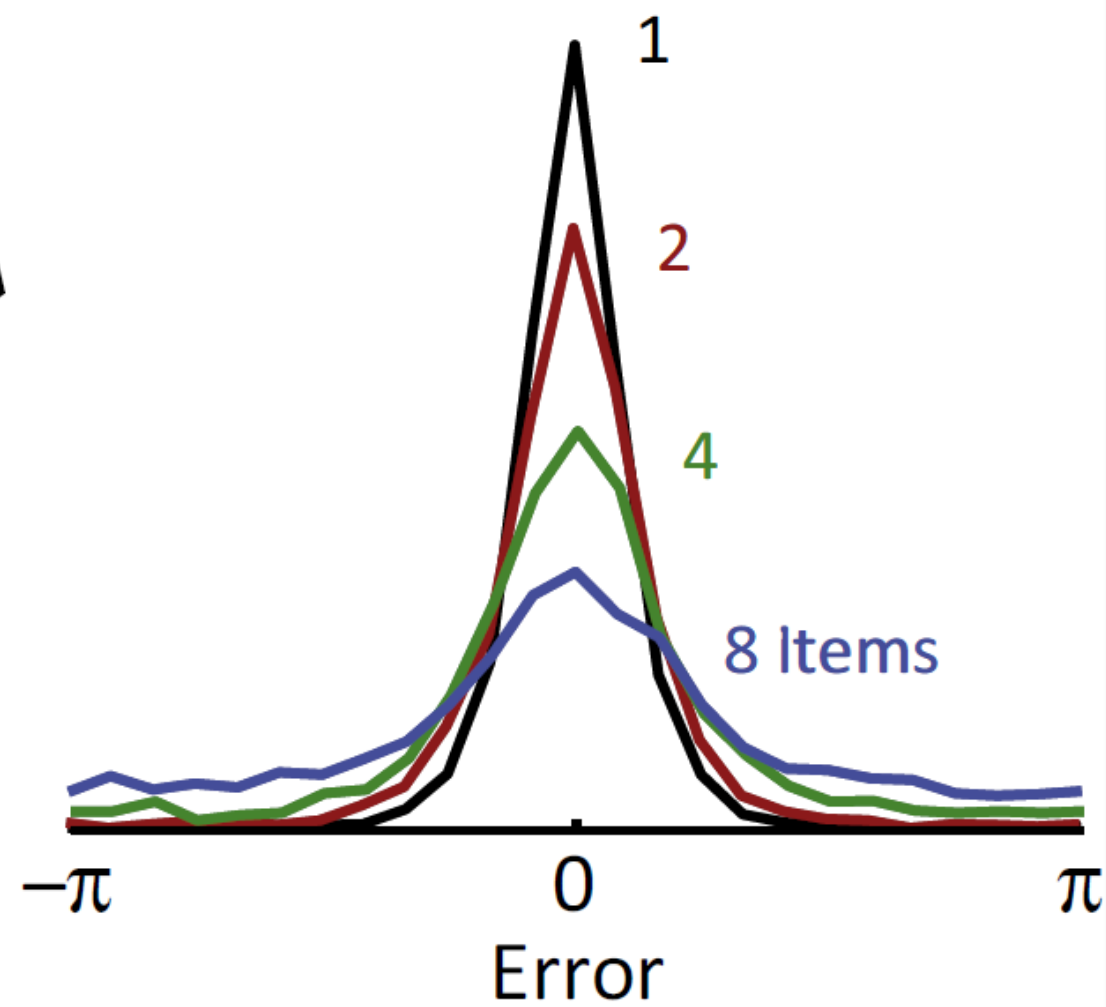
**Figure 1.** Errors in working memory (WM) reflect noise in neural coding. **(A)** An analogue report task, in which participants adjust a probe stimulus to match the orientation recalled at the same location in the memory array. **(B)** The distribution of responses around the true orientation changes as a function of the number of items in the memory array. **(C)** Error variability (black) increases monotonically with the number of array items. Curves show predictions of the population coding model (red; [28]) and slots + averaging model (blue; [27]) with best-fitting parameters. **(D)** Kurtosis (black) measures deviations from normality of error distributions. Curves as in (C). **(E)** The population coding model. Each stimulus is encoded by a set of orientation-selective neurons with bell-shaped (normal) tuning functions. Normalisation operates across the whole population, scaling summed activity to a fixed level. Neurons generate spikes according to a noisy Poisson process, and recall is modelled as maximum likelihood (ML) decoding of the spiking activity over a fixed time window. **(F)** Error distributions predicted by the population coding model with ML parameters (compare with B). Adapted and reprinted from [28], with permission from the Society for Neuroscience.

As set size increases in the response dial task [data shown for  $n = \{1, 2, 4, 8\}$ ], variability increases steadily. Accuracy degrades **gradually**, not abruptly as the slot model suggests.

(A)



(B)





VWM error distributions **do not match** the normal distribution—they have more kurtosis.

*Therefore, **assuming** the noise is normally distributed or indicative of “guessing” may be incorrect (p. 432).*

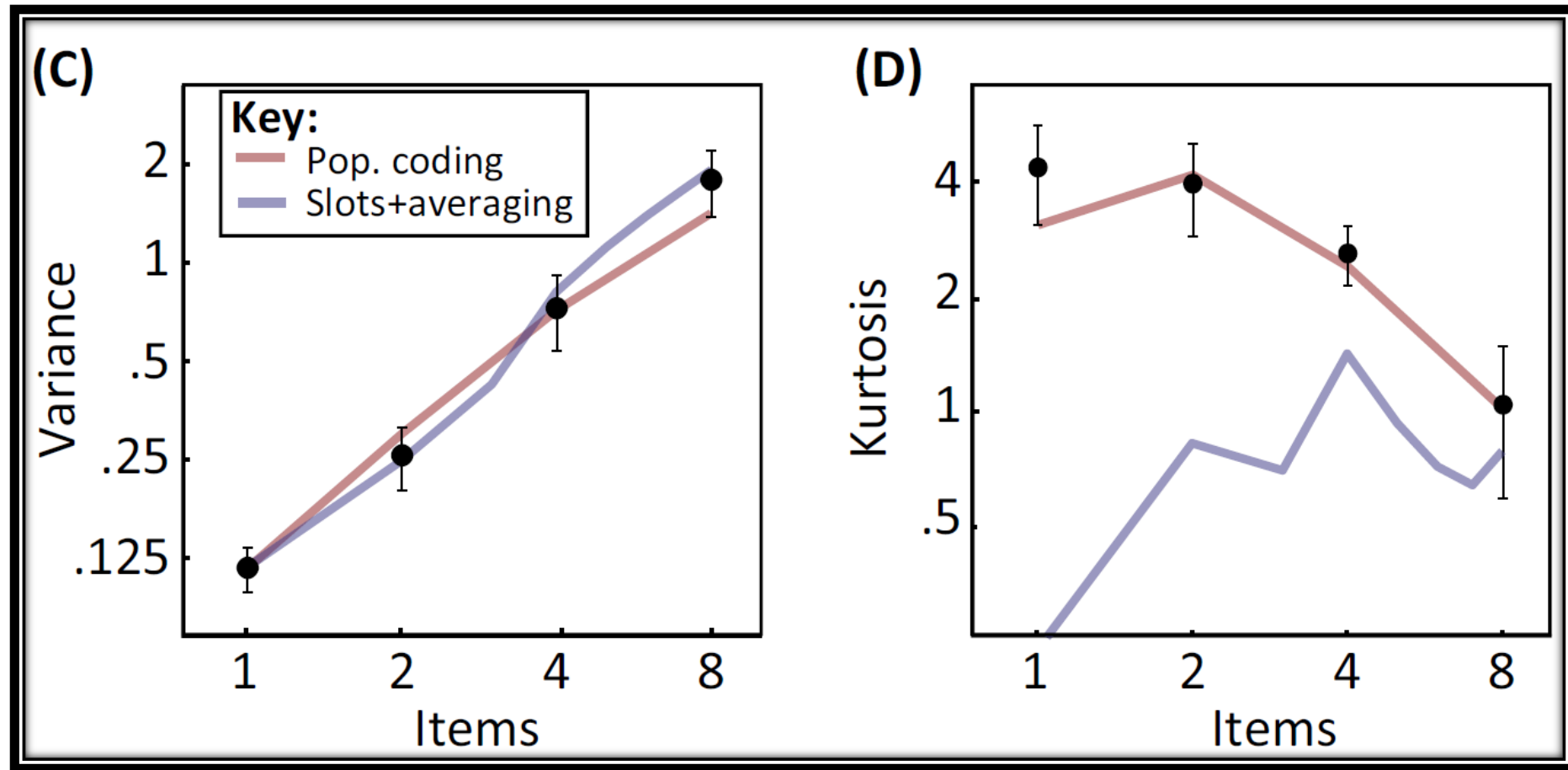


Figure 1-C: Log–log axes indicating that variance should increase monotonically with array size (p. 432).

Figure 1-D: Kurtosis from actual experiments is non-normal.

**Recall** that Brady, Konkle, & Alvarez (2011) argued slots are fungible (p. 4)—for instance, all the slots can be dedicated to one item to represent it with increased fidelity.

*Does **Bays** (2015) consider this?*

**Yes.**

Bays cites the “slots + averaging” model (p. 432–33), which proposes that 2 or more slots can contain independent representations of the **same** visual item. These slots are “averaged” to reconstruct the image more accurately.

Bays contends that, like the traditional slots model, the slots + averaging model **fails** to replicate the kurtosis found in actual data (p. 433), especially for a small number of items, including **one item**.

# Population Coding

A pool of neurons **shares** encoding of an item. “Common throughout the nervous system, including visual cortex” (p. 433) — **robust**, because any one neuron can fail with little impact.

*Redundancy — I think of this like a RAID 5 or RAID 6 array of hard disk drives.*

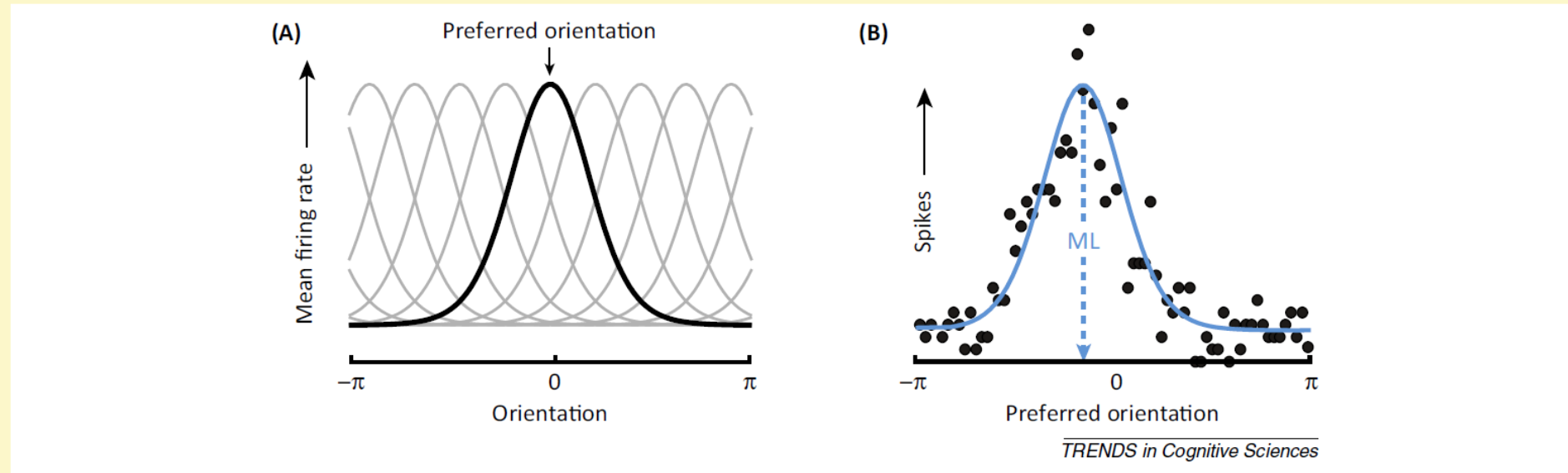
## Box 1. Population coding

Population coding is a method of encoding information in the combined activity of a pool of neurons [61,62]. The firing rate of each neuron is determined by a tuning function (Figure 1A) with a peak at a 'preferred' value that differs from neuron to neuron. This coding strategy is common throughout the nervous system, including visual cortex, where the encoded values are stimulus features such as orientation, and motor cortex, where values correspond to motor outputs such as an intended movement direction [63,64].

Population coding has the advantage that it is robust: because the encoded information is distributed over many neurons, damage to any one neuron has relatively little effect on the representation. Theoretical schemes have been proposed whereby population codes can store multiple inputs and represent uncertainty in the input [65,66]. Methods also exist by which information in a population code can persist over time (Box 2), and persistent activity associated

with WM in cortex, for example in prefrontal neurons, exhibits population coding [67].

Neural firing is probabilistic, thus information encoded by neurons can in general only be recovered imperfectly; because outputs can be averaged over many neurons, population codes help reduce this uncertainty. A particularly important method of decoding, which is in many situations statistically optimal, is maximum likelihood (ML). ML decoding [68] consists of identifying the represented value for which the observed activity was most likely to occur. This is also the value at the peak of an idealised (noiseless) response function that best fits the observed firing rates (Figure 1B). Biologically plausible methods have been identified for obtaining ML estimates from population codes [69,70]. However, it is unclear whether decoding takes place explicitly in the brain, or is only implicit in the transformation between variables, for example, from sensory to motor coordinates [71].



**Figure 1.** Standard model of population coding. **(A)** The mean activity of a neuron in response to an input (e.g., the orientation of a visual stimulus) is determined by its tuning function, which is typically bell shaped. Neurons within a population differ in their preferred stimulus value, that is, the value that elicits their maximum firing rate. **(B)** Population activity (black filled dots) plotted against the preferred stimulus value of each neuron. In any finite time period, the spike rate will be a noisy approximation to the idealised value set by the tuning function. Maximum likelihood finds the idealised output (blue curve) that most closely fits the actual activity [61].

# What does population coding do?

It limits spiking via **normalization** and distribution among visual items, giving a “plausible biological basis” for VWM as a limited resource (p. 432).

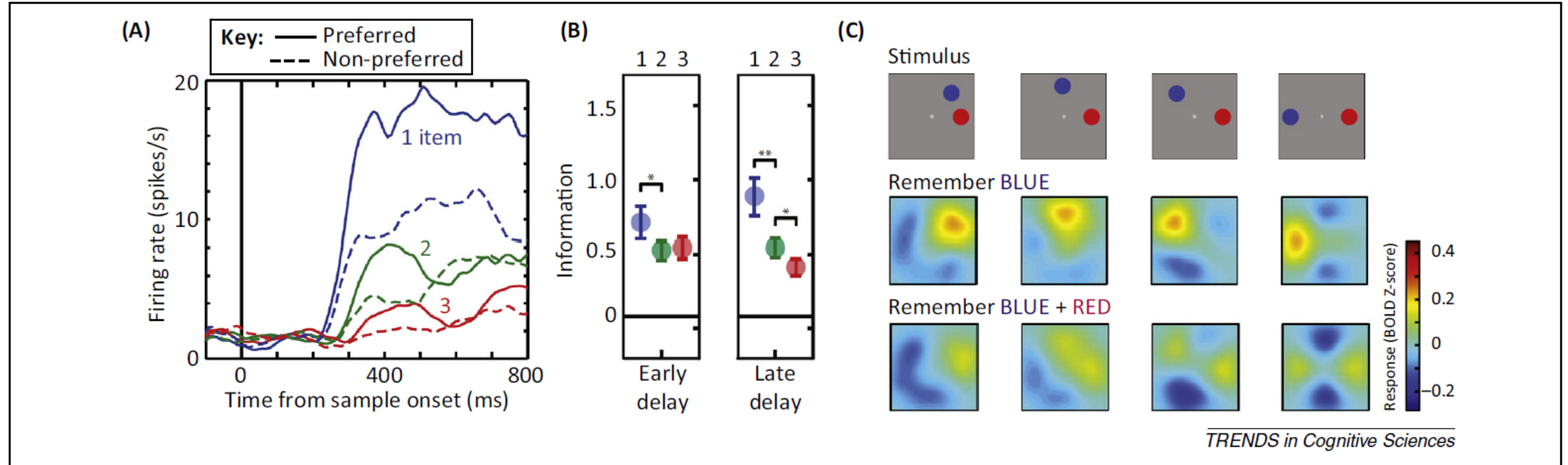


**Population coding** is provided as neurophysiological evidence to **support** the author's position, as is normalization, diffusion, and accumulation to bound (p. 437).

## Normalization (p. 433–34)

“Explains why variability increases with the number of items” (p. 433).

New fMRI evidence suggests this is a **broad** phenomenon that occurs across many stimuli at once, and even across multiple brain regions (p. 434).



**Figure 2.** Evidence for normalisation in working memory (WM)-related neural activity. **(A)** Firing rate of an example prefrontal neuron with persistent WM activity. Firing rate declines with increasing memory load, whether the stimulus in the receptive field corresponds to a preferred (unbroken lines) or non-preferred (broken lines) feature of the neuron. **(B)** Correspondingly, the information about a stimulus that can be extracted from delay-period neural activity declines as the number of items increases, consistent with normalisation. (A,B) Adapted and reprinted from [22], with permission from the National Academy of Sciences. Data shown are for lateral prefrontal cortex (LPFC); similar results were found in frontal eye field (FEF) and lateral intraparietal (LIP) areas. Note that significant differences between two and three items were observed only in the late phase of the delay in the LPFC; however, this was not a consistent observation across brain areas. **(C)** An fMRI study [24] decoded signals recorded during a spatial WM task (top row). When participants were instructed to remember the location of one of two stimuli, strong delay-period activity was observed corresponding to the remembered stimulus location only (middle row). When both stimuli were to be remembered, activity was observed corresponding to both locations, but with significantly reduced amplitude at each. Adapted and reprinted from [24], with permission from Elsevier. Data shown are from area V4A; similar results were found throughout visual and posterior parietal cortex.

## Decay (p. 434–35)

- VWM items become less accurate the longer they are maintained.
- More items to remember => faster decay
- “Cueing” an item helps to preserve it, but other items decay **faster**

# The Attractor Model (p. 434–35)

A possible neurophysiological explanation for decay:

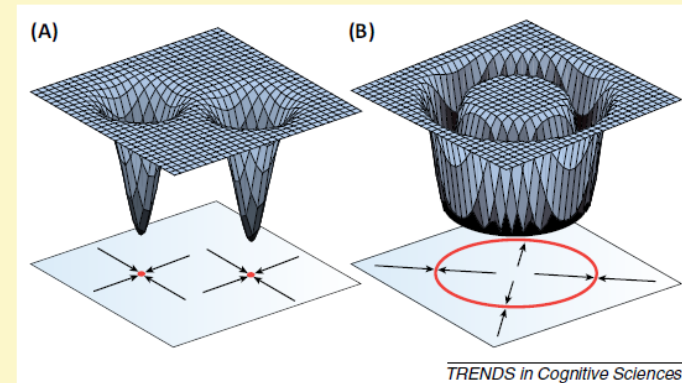
- A neural circuit that **sustains** patterns
- It seems it **diffuses** over time, rather than declining in amplitude

## Box 2. Attractor models

An attractor network [72,73] is a neural circuit for which certain patterns of activity are stable and self-sustaining. Regardless of how the network is initialised, activity will settle into one of these stable patterns or 'attractor states'. Such behaviour arises naturally in computational models of networks with recurrent excitation and is considered a possible basis for persistent activity underlying WM.

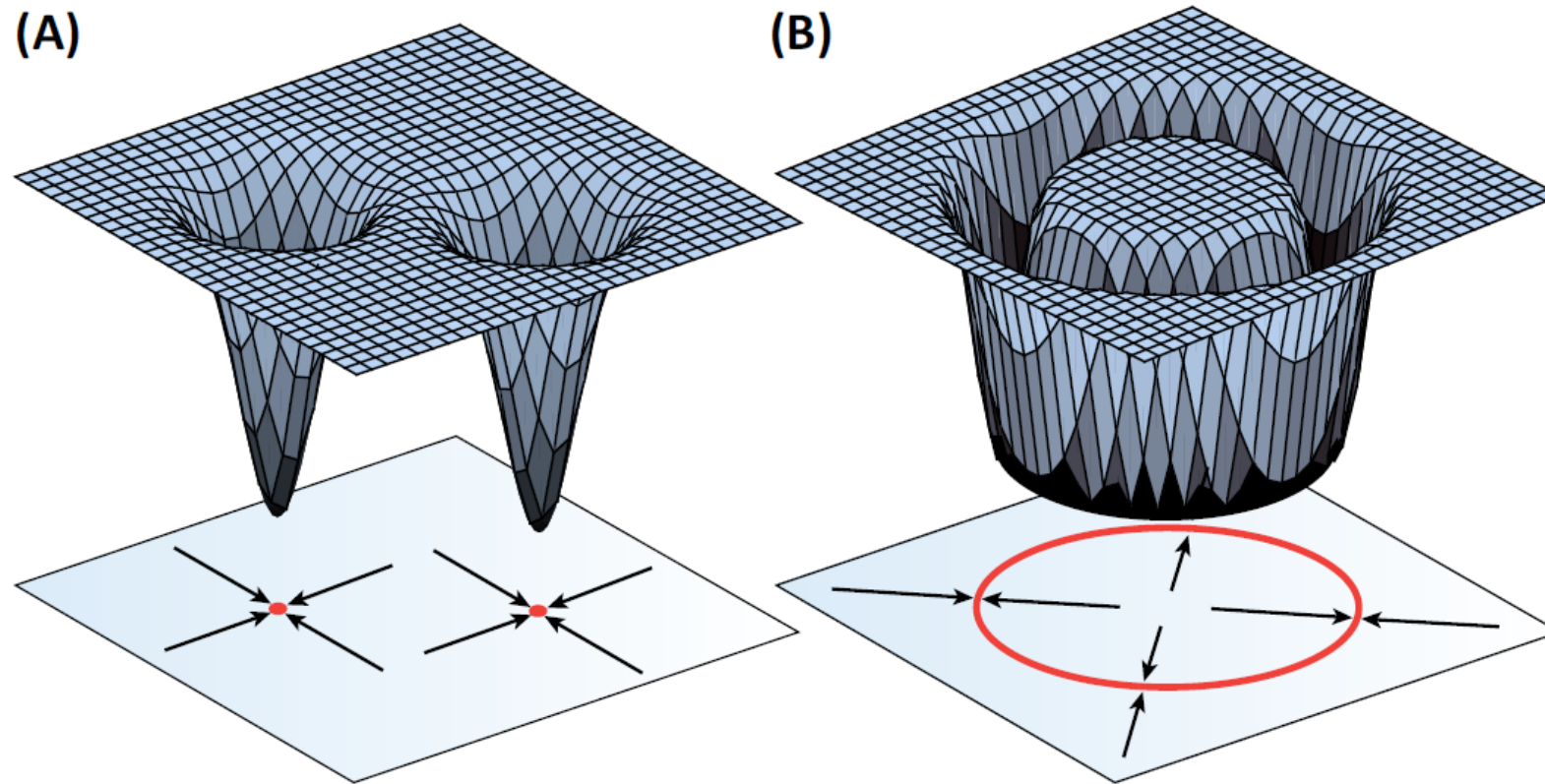
A discrete attractor network [74,75] has one or more fixed points of attraction (Figure 1A). Such a network can maintain the memory of a categorical variable, with each fixed point corresponding to a different discrete value (e.g., a letter of the alphabet). Discrete attractor networks can carry out a type of pattern completion: if the network is initialised with an input that does not correspond exactly to one of the categories, the activity will shift to the nearest stable state corresponding to the closest-matching category.

Continuous attractor models [76–78] extend this principle to networks with a continuum of attractor states. For example, memory for a continuous variable such as orientation can be stored in a 'ring attractor' network (Figure 1B), where every possible orientation corresponds to a different point on the ring of stable states. Similar to discrete attractors, these networks are relatively resistant to internal noise, in that the system rapidly corrects for perturbations away from the attractor line. However, perturbations along the attractor are not corrected, making the network susceptible to random drift ('diffusion') in its stored variable.



*TRENDS in Cognitive Sciences*

**Figure 1.** Attractor networks. (A) Energy landscape and state-space representation of a discrete attractor network with two fixed points (red dots). Here, two state variables ( $x$  and  $y$  axes) stand in for the high-dimensional state of activity in the network. Activity evolves over time (arrows) to settle in the nearest of the fixed attractor states. (B) Representation of a ring attractor network. Activity is stable anywhere on the continuum of attractor states (red line).



*TRENDS in Cognitive Sciences*

**Figure I.** Attractor networks. **(A)** Energy landscape and state-space representation of a discrete attractor network with two fixed points (red dots). Here, two state variables ( $x$  and  $y$  axes) stand in for the high-dimensional state of activity in the network. Activity evolves over time (arrows) to settle in the nearest of the fixed attractor states. **(B)** Representation of a ring attractor network. Activity is stable anywhere on the continuum of attractor states (red line).

## The Attractor Model (cont.)

*This is the main issue that Bays identifies with using this as a model of VWM:*

The normalized attractor model does not work with **analog** recall tasks such as recalling two similar colors; two similar stimuli simply merge in this model (p. 435).

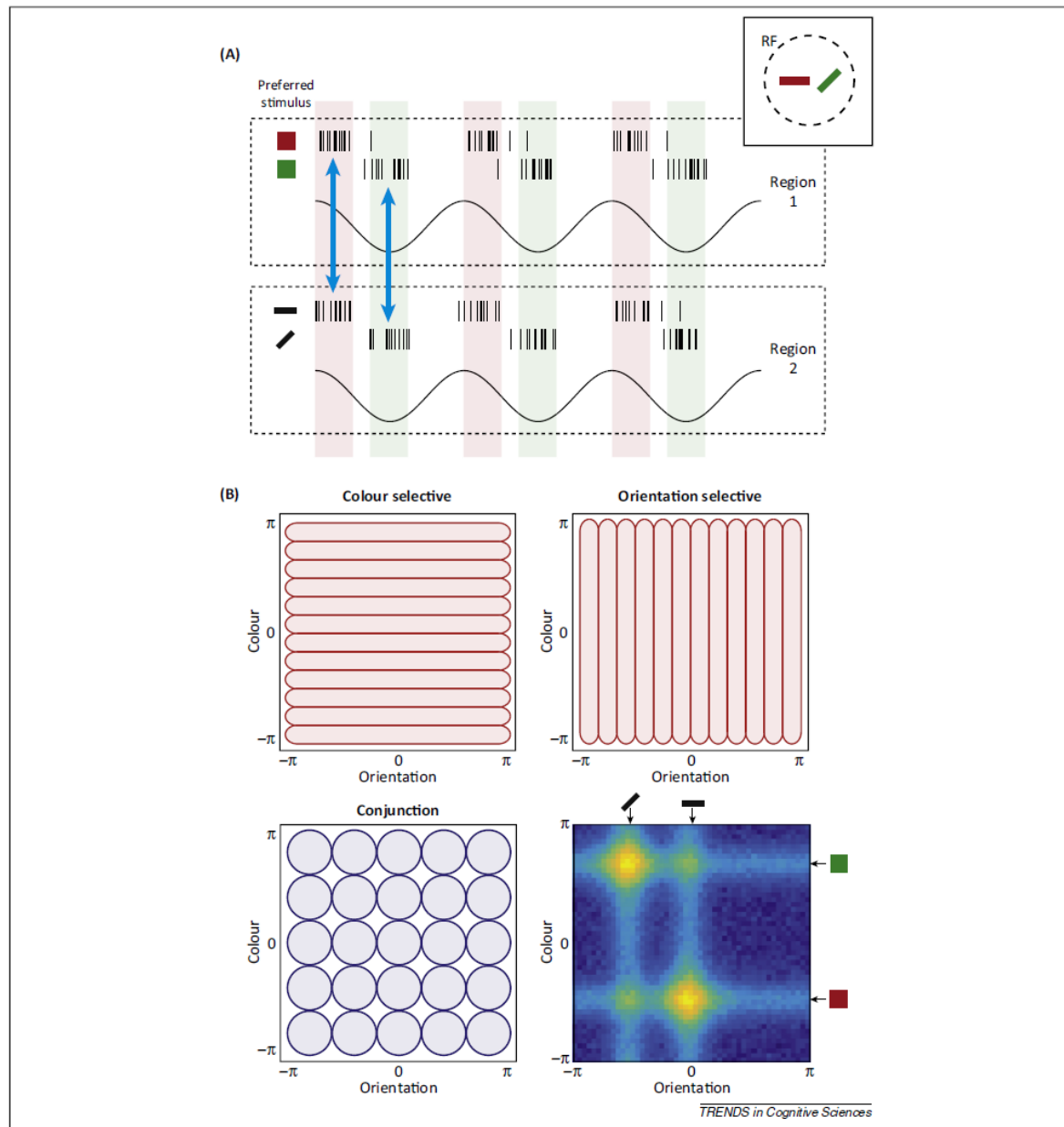


## Recall Latency (p. 435)

- As the number of VWM items increases, latency increases
  - A strongly skewed distribution
  - Decay continues even during retrieval
- Like an **accumulation** process—reaches a “threshold” where the stimulus can be retrieved (p. 435).

## **Binding Errors (p. 435–37)**

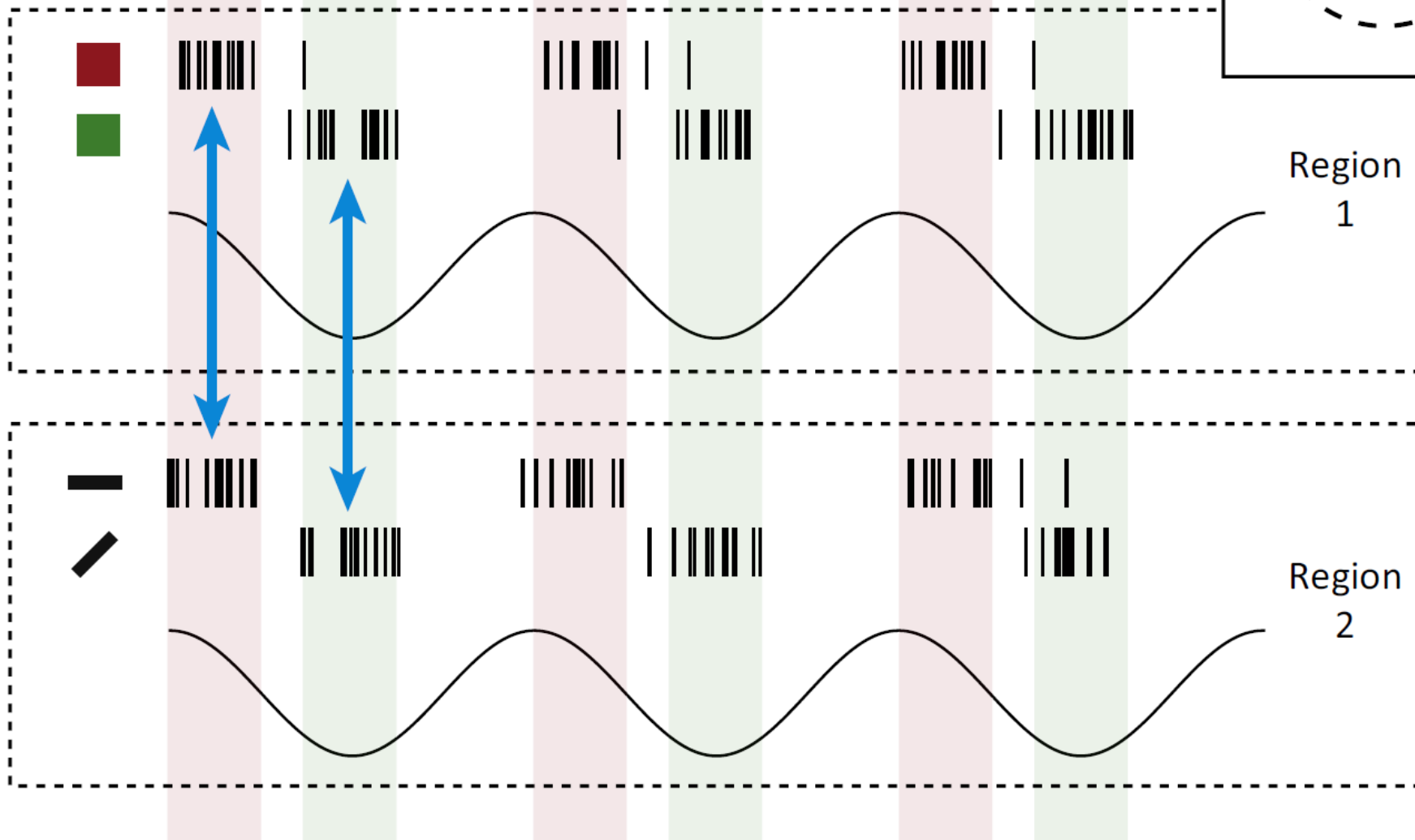
- Occur when visual features are bound to the wrong objects
- Result in inaccurate recall of what was seen
- Uncommon in perception; common in VWM
- Might arise because spike timing is stochastic



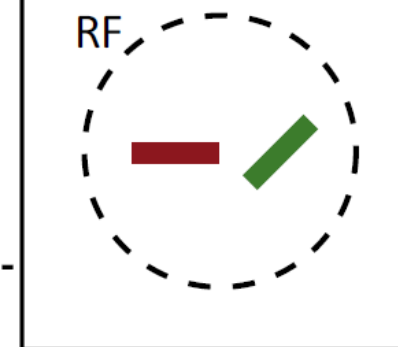
**Figure 3.** Models of binding and misbinding in working memory (WM). (A) Binding by synchrony. Neurons in two brain regions synchronise their activity to a common waveform (e.g., theta rhythm) in order to store in memory (inset) a red horizontal bar and a green oblique bar. A red-selective neuron in colour region 1 fires spikes in the same phase as a horizontal-selective neuron in orientation region 2; likewise, activity of a green-selective neuron is matched with an oblique-selective neuron. Spike timing is probabilistic, and errors of binding (e.g., recall of a green horizontal bar) arise from mistimed spikes. (B) Binding by conjunction. Stimulus information is stored in the activity of a mixed population of neurons, comprising colour-selective neurons (response fields shown top-left), orientation-selective neurons (top-right), and conjunction neurons that are activated only by a specific pairing of colour and orientation (bottom-left). Optimal decoding of the population activity recovers the stimulus values and their binding (decoding probability shown bottom-right). Misbinding occurs probabilistically because of noise in firing rate.

(A)

Preferred  
stimulus



RF



## Binding Errors (cont.)

Bays' argument: Because binding errors can **only** occur between items in memory, *if* there is a “hard” limit on VWM like slot models propose, *then* binding errors should reach a **plateau** once that limit is exceeded.

*However, binding errors continue to **increase**.*

# Overview

Bays overall argument, mentioned in the abstract, is that VWM is a **continuous** resource that degrades gracefully, rather than a discrete resource that degrades spectacularly.

*Similar to an analog versus digital dichotomy*

## Overview (cont.)

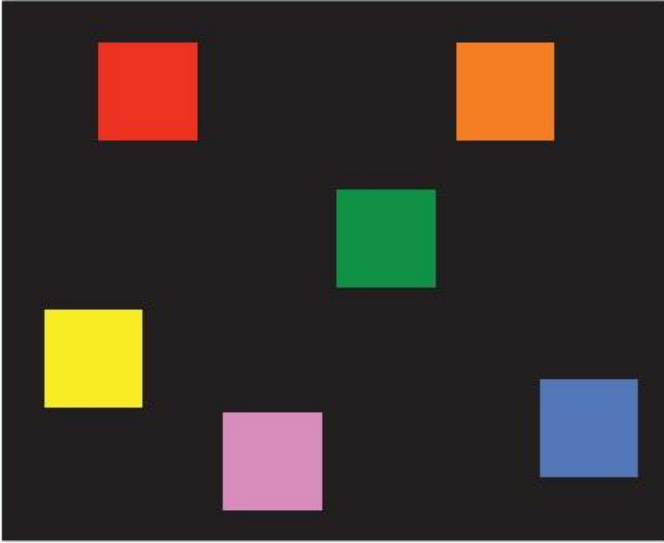
“Currently, no model incorporating a deterministic limit has been shown to reproduce the characteristic deviations from normality observed in [VWM] errors, and this is an important challenge for proponents of this view” (p. 433).

# Discussion

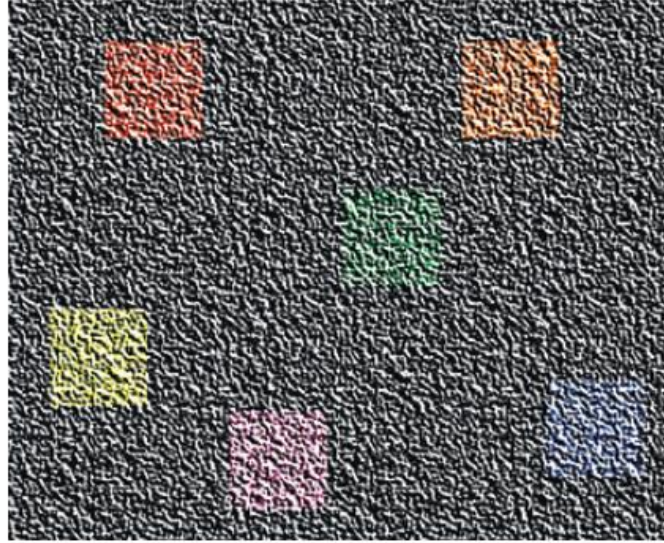
Luck & Vogel (2013) reference a study finding that subjects cannot “trade precision for capacity” even when money was offered (p. 396)!



(A) You see this

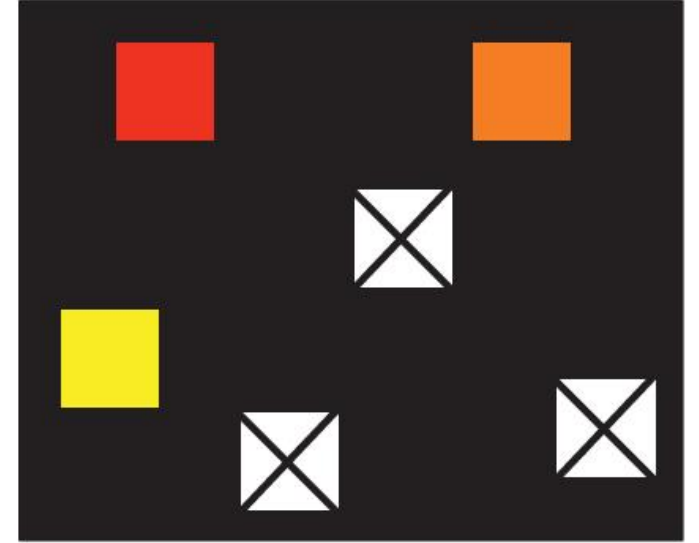


Do you remember this?



Continuous resource

Or this?



Discrete slots

Luck & Vogel (2013) provide this figure to help visualize the arguments (p. 394).

## Discussion (cont.)

Luck & Vogel (2013) do not address Bays' (2015) kurtosis / abnormality argument, but a response may be forthcoming.

*Is kurtosis the foundation for Bays' argument?*

*If so, is it a **weak** foundation?*

*Is this a loaded question?*

## Discussion (cont.)

What do you think? Is visual working memory best characterized by a slot model? Perhaps there should just be more slots (i.e. 6 instead of 3–4)?

*Is the resource / stochastic model superior, as Bay contends?*



Your PC ran into a problem and needs to restart. We're just collecting some error info, and then we'll restart for you. (0% complete)

If you'd like to know more, you can search online later for this error: HAL\_INITIALIZATION\_FAILED

Oh no! I ran out of slots!

## Discussion (cont.)

Is Bays being biased?

What about Luck & Vogel?

Is this factionalism (or partisanship)?

If so, is it **aiding** or **hindering** scientific progress in this area?

## Discussion (cont.)

Who thinks a more accurate model may be a **mix** of both models?

Which elements from each model might be **supported** or **unsupported**?

### Box 3. Outstanding questions

---

- How can the stochastic view of WM be extended to memoranda that are categorical in nature, for example, letters, shapes, or high-level objects? Can discrete attractor networks ([Box 2](#)) provide a useful description?
- How are time intervals and the temporal order of events represented in WM?
- Can the population coding model of visual WM be adapted for other modalities, for example, tactile or auditory WM?
- Are there limits on flexible allocation? A system based on binary spiking events could be practically limited in how finely activity can be distributed between representations. Furthermore, under certain conditions it may be statistically optimal to limit allocation to a subset of available items, rather than store all with very low precision. Such effects might be captured by a stochastic upper bound, as in [\[54\]](#).
- What is the neurophysiological basis for the representation in WM of ensemble statistics [\[79\]](#), such as the average colour of a group of objects?
- What is the mechanism underlying the broad normalisation observed in WM-related activity, and how is it distinct from those proposed for perception and attention [\[33\]](#)? Possible mechanisms include lateral inhibition within posterior areas, or a resource-limited descending signal from, for example, prefrontal cortex.

In conclusion, Bays concedes that the connections between behavioral observations and neurophysiology are **speculative** and theoretical—further research is required (p. 437).



# References

- Bays, P. M. (2015). Spikes not slots: Noise in neural populations limits working memory. *Trends in Cognitive Sciences*, 19(8), 431–438.  
<http://dx.doi.org/10.1016/j.tics.2015.06.004>
- Brady, T., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, 11(5), 1–34.  
doi:10.1167/11.5.4
- Luck, S. J. & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to visual differences. *Trends in Cognitive Sciences*, 17(8), 391–400.  
<http://dx.doi.org/10.1016/j.tics.2013.06.006>

# References

*Figures were primarily from the Bays (2015) article.*

*The conceptual figure with colored squares for “continuous resource” versus “discrete slots” was from the Luck & Vogel (2013) article.*

*The Super Mario 64 screenshot, analog television image, and Windows “blue screen of death” screenshot were found via Google Image Search. Images in this PowerPoint presentation are hyperlinks to the source webpages.*